Journal of Multivariate Analysis 102 (2011) 1472–1488

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Testing model assumptions in functional regression models

Axel Bücher, Holger Dette *, Gabriele Wieczorek

Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany

ARTICLE INFO

Article history:
Received 22 October 2009
Available online 30 May 2011

AMS 2000 subject classification:
62G10

Keywords:
Goodness-of-fit tests
Functional data
Parametric bootstrap
Tests for heteroscedasticity

ABSTRACT

In the functional regression model where the responses are curves, new tests for the functional form of the regression and the variance function are proposed, which are based on a stochastic process estimating L^2-distances. Our approach avoids the explicit estimation of the functional regression and it is shown that normalized versions of the proposed test statistics converge weakly. The finite sample properties of the tests are illustrated by means of a small simulation study. It is also demonstrated that for small samples, bootstrap versions of the tests improve the quality of the approximation of the nominal level.

1. Introduction

Since the pioneering work by Ramasay and Dalzell [25] on regression analysis for functional data, this topic has received considerable attention in the recent literature. The interest in statistical techniques, enabling us to take into account the functional nature of data, stems from the fact that nowadays in many applications (for instance in climatology, remote sensing, linguistics etc.), the data comes from the observation of a continuous phenomenon over time. For a review on the statistical problems and techniques for functional data, we refer to the monographs of [26,16]. In these models either predictors or responses can be viewed as random functions. The data typically appears when the value of a variable is repeatedly recorded on a dense grid of time points for a sample of subjects. While many authors consider the problem of estimating the regression or generalizing classical concepts of multivariate statistics as the principal component or discrimination analysis to the situation where the data are curves (see for example [2,14,5,20,9,15,13] or [23] among many others), much less attention has been paid to the problem of testing model assumptions when analyzing functional data.

Many authors discussed the problem of testing hypotheses in a linear functional data model. For example, Cardot et al. [7], Müller and Stadtmüller [23] and Cardot et al. [8] considered the problem of testing a simple hypothesis in the case where the response is real and the predictor is a random function, while Mas [22] investigated a test for the mean of random curves. Recently, Shen and Faraway [29] and Yang et al. [31] discussed an F-test in a linear longitudinal data model, while Kokoszka et al. [21] tested for lack of dependence in a functional linear model where both the response and the predictor are curves. We also refer to the recent Ph.D. thesis of Delsol [10] where a test is constructed on the basis of kernel estimates.

The present work considers the problem of testing model assumptions in the nonparametric functional regression model

$$Y_{i,n}(u) = m(u, t_{i,n}) + \varepsilon(u, t_{i,n}), \quad t_{i,n} \in [0, 1],\ i = 1, \dots, n, \tag{1.1}$$

where u varies (without loss of generality) in the interval $[0, 1]$. Our main concern deals with the problem of validating a parametric assumption of the form

$$Y_{i,n}(u) = g(u, t_{i,n}, \beta) + \varepsilon(u, t_{i,n}), \quad t_{i,n} \in [0, 1],\ i = 1, \dots, n, \tag{1.2}$$

* Corresponding author.
E-mail addresses: axel.buecher@ruhr-uni-bochum.de (A. Bücher), holger.dette@ruhr-uni-bochum.de, holger.dette@rub.de (H. Dette).

0047-259X/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2011.05.014

where g is a given parametric functional regression model and $\beta : [0, 1] \to \mathbb{R}^k$ denotes a function, which depends either on the variable u or on t (note that both cases correspond to a different parametric modeling of the functional data). The latter model has been considered in the linear context by numerous authors. In particular Shen and Faraway [29] and Yang et al. [31] have proposed generalizations of the F-test for model $Y(u) = x^T\beta(u) + \varepsilon(u)$ and used these methods to analyze some data from Ergonomics. While in this work the predictor x is considered as discrete (as in the classical ANOVA model) we concentrate in the present paper on the case where the variable t in (1.2) varies in continuous way. Our work is inspired by the recent paper of Hlubinka and Prchal [19], who proposed a functional regression model of the form (1.2) to study the time-variation of vertical atmospheric radiation profiles. These authors assumed that the parameter (function) β depends on the value t.

In Section 2 we introduce some notation and propose a test for the hypothesis that the regression function in the nonparametric functional regression model (1.1) is of a specific parametric form as given in (1.2) with a function β depending on the variable u, that is

$$H_0 : m(u, t) = g(u, t, \beta(u)) \tag{1.3}$$

for some parametric function g and a parameter $\beta : [0, 1] \to \mathbb{R}^k$. The case, where the parameter depends on t is investigated in Section 3, where we consider the hypothesis

$$H_0 : m(u, t) = h(u, t, \gamma(t)) \tag{1.4}$$

for a parametric function h and some function $\gamma : [0, 1] \to \mathbb{R}^k$. Finally, we discuss in Section 4 the problem of testing parametric assumptions regarding the second order properties of the process $Y(u)$. More precisely, if $r(t, u, v) = \mathrm{Cov}(\varepsilon(u, t), \varepsilon(v, t))$ denotes the covariance of the observations $Y(u)$ and $Y(v)$, we are interested in the hypothesis

$$H_0 : r(t, u, v) = r(u, v), \tag{1.5}$$

which corresponds to the case of homoscedasticity. Note that this assumption is necessary for the application of the F-tests proposed by Shen and Faraway [29] and Yang et al. [31]. Moreover, this assumption was also made by Hlubinka and Prchal [19] who proposed a nonlinear functional regression model for the analysis of changes in atmospheric radiation. The proposed tests for the hypotheses (1.3)–(1.5) are very simple and are based on stochastic processes of empirical L^2-distances between the nonparametric and parametric functional regression model. We prove weak convergence of these processes under the null hypothesis and fixed alternatives and, as a consequence, asymptotic normality of functionals of these processes. In Section 5 we demonstrate by means of a simulation study that for moderate sample sizes the quantiles of the asymptotic distribution provide a rather accurate approximation of the nominal level. On the other hand – for small sample sizes – a wild bootstrap version of the test is proposed and its accuracy is also investigated. Finally, some technical details are given in an Appendix.

2. A process of empirical L^2-distances for testing (1.3)

Consider the nonparametric functional regression model defined by (1.1) and assume that a triangular array of n independent observations $\{Y_{i,n} \mid i = 1, \ldots, n\}$ is available at distinctive points $0 \le t_{1,n} < \cdots < t_{n,n} \le 1$. In order to keep the notation as simple as possible we omit the second index and write Y_i for $Y_{i,n}$, t_i for $t_{i,n}$ etc. For the discussion of the asymptotic properties of the tests proposed in this paper we will assume that the design points $t_1, \ldots, t_n$ satisfy

$$\max_{i=2}^{n} \left| \int_{t_{i-1}}^{t_i} h(t)\,\mathrm{d}t - \frac{1}{n} \right| = o(n^{-(1+\gamma)}), \tag{2.1}$$

where $h \in \mathrm{Lip}_\gamma[0, 1]$ is a strictly positive (unknown) density on the interval $[0, 1]$, which is Lipschitz continuous of order $\gamma > 1/2$ (see [27]). Note that the choice $h(t) \equiv 1$ corresponds to an asymptotically uniform design. For example, if $t_i = \frac{i}{n+1}$ the error is in fact of order $O(n^{-2})$. Similarly other designs could be considered by choosing an appropriate limiting density h. For example the function $h(x) = 2x$ corresponds to a design which takes more observation in the right part of the interval $[0, 1]$, that is $t_i = \sqrt{\frac{i}{n+1}}$ $(i = 1, \ldots, n)$.

For the construction of a test for the hypothesis (1.3) of a parametric functional regression model we consider the class of parametric models

$$\mathcal{M} = \{g(\cdot, \cdot, \beta(\cdot)) : [0, 1] \times [0, 1] \longrightarrow \mathbb{R} \mid \beta : [0, 1] \longrightarrow \Theta\},$$

where Θ is some subset of $\mathbb{R}^k$. For the sake of transparency we first discuss the case of testing the hypothesis of a linear functional regression model, that is

$$H_0 : m(u, t) = g(u, t, \beta(u)) = \beta(u)^T f(u, t), \tag{2.2}$$

where $f(u, t) = (f_1(u, t), \ldots, f_k(u, t))^T$ are given regression functions. We define for fixed $u \in [0, 1]$ the inner product

$$\langle p, q \rangle_u = \int p(u, t) q(u, t) h(t)\,\mathrm{d}t.$$

On the space of functions defined on the unit square $[0, 1]^2$ with corresponding norm $\|\cdot\|_u$, and consider

$$M_u^2 = \inf_{\beta(u)} \|m(u, \cdot) - \beta(u)^T f(u, \cdot)\|_u^2$$

as the minimal distance from m to functions of the form (2.2). A standard result from Hilbert space theory (see [1]) yields that M_u^2 can be expressed as a ratio of two Gramian determinants, i.e.

$$M_u^2 = \frac{\Gamma_u(m, f_1, \ldots, f_k)}{\Gamma_u(f_1, \ldots, f_k)},$$

where $\Gamma_u(p_1, \ldots, p_k) = \det(\langle p_i, p_j\rangle_u)_{i,j=1,\ldots,k}$ is the Gramian determinant of the function $p_1, \ldots, p_k$. In order to obtain an estimator for M_u^2 we replace the inner products $A_{u,0} = \langle m, m\rangle_u$, $A_{u,p} = \langle m, f_p\rangle_u$ and $B_{u,p,q} = \langle f_p, f_q\rangle_u$ by their empirical counterparts

$$\hat{A}_{u,0} = \frac{1}{n-1}\sum_{i=2}^{n} Y_i(u)Y_{i-1}(u),$$

$$\hat{A}_{u,p} = \frac{1}{n}\sum_{i=1}^{n} Y_i(u) f_p(u, t_i),$$

$$\hat{B}_{u,p,q} = \frac{1}{n}\sum_{i=1}^{n} f_p(u, t_i) f_q(u, t_i),$$

where $p, q = 1, \ldots, k$. Note that we estimate $A_{u,0}$ by $\frac{1}{n-1}\sum_{i=2}^{n} Y_i(u)Y_{i-1}(u)$ instead of $\frac{1}{n}\sum_{i=1}^{n} Y_i^2(u)$ since the latter estimator is asymptotically biased. This yields a canonical estimate

$$\hat{M}_u^2 = \frac{\begin{vmatrix} \hat{A}_{u,0} & \hat{A}_{u,1} & \cdots & \hat{A}_{u,k} \\ \hat{A}_{u,1} & \hat{B}_{u,1,1} & \cdots & \hat{B}_{u,1,k} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{A}_{u,k} & \hat{B}_{u,k,1} & \cdots & \hat{B}_{u,k,k} \end{vmatrix}}{\begin{vmatrix} \hat{B}_{u,1,1} & \cdots & \hat{B}_{u,1,k} \\ \vdots & \ddots & \vdots \\ \hat{B}_{u,k,1} & \cdots & \hat{B}_{u,k,k} \end{vmatrix}} \tag{2.3}$$

of the L^2-distance M_u^2. In the following discussion we will study the asymptotic properties of the process $\{\hat{M}_u^2\}_{u\in[0,1]}$. Denote by

$$\mathrm{Lip}_\gamma^{unif}[0, 1] = \{f = f(x, \cdot) : |f(x, t) - f(x, s)| \le C|s - t|^\gamma; s, t \in [0, 1]\} \tag{2.4}$$

the set of all functions $f : [0, 1] \times [0, 1] \to \mathbb{R}$ satisfying a uniform Lipschitz condition (in other words, the constant C in (2.4) does not depend on x) and assume that for some $\gamma > 1/2$ and for all $(t, u, v) \in [0, 1]^3$

$$f_j(u, \cdot), f_j(\cdot, t) \in \mathrm{Lip}_\gamma^{unif}[0, 1] \quad j = 1, \ldots, k$$

$$m(u, \cdot), m(\cdot, t) \in \mathrm{Lip}_\gamma^{unif}[0, 1],$$

$$r(\cdot, u, v), r(t, \cdot, v), r(t, u, \cdot) \in \mathrm{Lip}_\gamma^{unif}[0, 1],$$

where

$$r(t, u, v) = E[\varepsilon(u, t)\varepsilon(v, t)]$$

denotes the covariance of the (centered) errors at the point t. We also assume that the processes $\{\epsilon(u, t_i)\}_{u\in[0,1]}$ and $\{\epsilon(u, t_j)\}_{u\in[0,1]}$ are independent whenever $i \neq j$. The following result specifies the asymptotic properties of the stochastic process $\{\sqrt{n}(\hat{M}_u^2 - M_u^2)\}_{u\in[0,1]}$. Throughout this paper the symbol $\Longrightarrow$ denotes weak convergence.

Theorem 2.1. *If the assumptions stated in this section are satisfied and the linear hypothesis (2.2) has to be tested we have as* $n \to \infty$

$$\{\sqrt{n}(\hat{M}_u^2 - M_u^2)\}_{u\in[0,1]} \Longrightarrow G,$$

in $C[0, 1]$, *where G is a centered Gaussian process with covariance $k(u, v)$ given by*

$$k(u, v) = \int r^2(t, u, v)h(t)\,\mathrm{d}t + 4\int r(t, u, v)(m(u, t) - g(u, t, \beta_0(u)))(m(v, t) - g(v, t, \beta_0(v)))h(t)\,\mathrm{d}t \tag{2.5}$$

and

$$\beta_0(u) = \operatorname{argmin}_\beta \|m(u,\cdot) - g(u,\cdot,\beta)\|_u^2 \tag{2.6}$$

corresponds to the parameter of the best approximation of the function $m(u,\cdot)$ by the parametric regression model.

Proof of Theorem 2.1. We assume without loss of generality that the functions $f_1(u,\cdot),\ldots,f_k(u,\cdot)$ are orthonormal with respect to the inner product $\langle p,q\rangle_u$. Then the minimal L^2-distance obtained by the best approximation simplifies to

$$M_u^2 = A_{u,0} - \sum_{p=1}^k A_{u,p}^2.$$

It is easy to see that the statistics $\hat{A}_{u,p}$ and $\hat{B}_{u,p,q}$ are $\sqrt{n}$ consistent estimates of the quantities $\langle m,f_p\rangle_u = \beta(u)$ and $\langle f_p,f_q\rangle = \delta_{p,q}$, respectively, and consequently we obtain for the statistic in (2.3)

$$M_n(u) := \sqrt{n}(\hat{M}_u^2 - M_u^2) = \sqrt{n}\left\{\hat{A}_{u,0} - \sum_{p=1}^k \hat{A}_{u,p}^2 - M_u^2\right\} + o_p(1) = \bar{M}_n(u) + o_P(1)$$

uniformly with respect to $u\in[0,1]$, where the last equality defines the process $\bar{M}_n(u)$ in an obvious manner. For the proof of weak convergence we have to show

$$(\bar{M}_n(u_1),\ldots,\bar{M}_n(u_m)) \xrightarrow{\mathcal{D}} (G(u_1),\ldots,G(u_m)) \quad \forall u_1,\ldots,u_m\in[0,1],\ m\in\mathbb{N}.$$

Tightness of the sequence $(\bar{M}_n)_{n\in\mathbb{N}}$.

The convergence of the finite dimensional distributions follows from Theorem 2.1 and its proof in [12]. For a proof of tightness we use the decomposition $\bar{M}_n(u) = U_n(u) + V_n(u)$ with

$$U_n(u) = \sqrt{n}\left(\hat{A}_{u,0} - E\hat{A}_{u,0} - \left(\sum_{p=1}^k \hat{A}_{u,p}^2 - E\hat{A}_{u,p}^2\right)\right)$$

$$V_n(u) = \sqrt{n}\left(E\hat{A}_{u,0} - \langle m,m\rangle_u - \left(\sum_{p=1}^k E\hat{A}_{u,p}^2 - \langle m,f_p\rangle_u\right)\right).$$

Consequently, it is sufficient to show that the (deterministic) sequence $V_n(u)$ converges uniformly to 0, i.e.

$$\sup_{u\in[0,1]} |V_n(u)| = o(1), \tag{2.7}$$

and that the process $\{U_n(u)\}_{u\in[0,1]}$ is tight. For this purpose we use Theorem 12.3 from [4] and show that there are constants $\alpha>0$, $\gamma\ge 0$ and a nondecreasing, continuous function F on $[0,1]$ such that

$$E\left[|U_n(u) - U_n(v)|^\gamma\right] \le |F(u) - F(v)|^\alpha. \tag{2.8}$$

We first prove (2.7) and introduce the decomposition

$$V_n(u) = \sqrt{n}\left(V_{n0}(u) - \sum_{p=1}^k V_{np}(u)\right)$$

with $V_{n0}(u) = E\hat{A}_{u,0} - \langle m,m\rangle_u$ and $V_{np}(u) = E\hat{A}_{u,p}^2 - \langle m,f_p\rangle_u^2$. Assertion (2.7) follows from

$$\sup_{u\in[0,1]} |V_{np}(u)| = o(n^{-\frac{1}{2}}), \quad p = 0,\ldots,k. \tag{2.9}$$

We consider exemplarily the first summand $V_{n0}(u)$, which can be represented as

$$V_{n0}(u) = A_1(u) + A_2(u) + A_3(u)$$

with

$$A_1(u) = \frac{1}{n-1}\sum_{i=1}^n m^2(u,t_i) - \langle m,m\rangle_u$$

$$A_2(u) = -\frac{1}{n-1}\sum_{i=2}^n m(u,t_i)(m(u,t_i) - m(u,t_{i-1}))$$

$$A_3(u) = -\frac{1}{n-1} m(u,t_1).$$

Using the fact that $m^2(u,\cdot) \in \mathrm{Lip}_\gamma^{unif}[0,1]$ and taking into account that $\max_{i=2}^n |t_i - t_{i-1}| = O(n^{-\gamma}) = o(n^{-\frac{1}{2}})$, by (2.1), we obtain that all terms are of order $o(n^{-1/2})$, uniformly with respect to $u \in [0,1]$. This proves (2.9) for $p = 0$ and similar arguments for the remaining terms show that (2.7) holds.

In order to show that condition (2.8) is valid we calculate

$$E\left[(U_n(u) - U_n(v))^2\right] = n\,(B_1 + B_2 + B_3 + B_4),$$

where

$$B_1 = \mathrm{Var}(\hat{A}_{u,0}) + \mathrm{Var}(\hat{A}_{v,0}) - 2\,\mathrm{Cov}(\hat{A}_{u,0}, \hat{A}_{v,0}),$$

$$B_2 = \mathrm{Var}\left(\sum_{p=1}^k \hat{A}_{u,p}^2\right) + \mathrm{Var}\left(\sum_{p=1}^k \hat{A}_{v,p}^2\right) - 2\,\mathrm{Cov}\left(\sum_{p=1}^k \hat{A}_{u,p}^2, \sum_{p=1}^k \hat{A}_{v,p}^2\right),$$

$$B_3 = 2\,\mathrm{Cov}\left(\hat{A}_{u,0}, \sum_{p=1}^k \hat{A}_{v,p}^2\right) - 2\,\mathrm{Cov}\left(\hat{A}_{u,0}, \sum_{p=1}^k \hat{A}_{u,p}^2\right),$$

$$B_4 = 2\,\mathrm{Cov}\left(\sum_{p=1}^k \hat{A}_{u,p}^2, \hat{A}_{v,0}\right) - 2\,\mathrm{Cov}\left(\hat{A}_{v,0}, \sum_{p=1}^k \hat{A}_{v,p}^2\right).$$

We now show that it is possible to find, for each term $B_i = B_i(u,v)$ $(i = 1, \ldots, 4)$ a constant C such that

$$n\,B_i(u,v) \le C\,|u-v|^\gamma,$$

which proves condition (2.8). For this purpose we exemplarily consider the expression B_1, the corresponding statements for the other terms follow along similar lines. A straightforward but tedious calculation yields

$$\begin{aligned}\mathrm{Cov}(\hat{A}_{u,0}, \hat{A}_{v,0}) &= \frac{1}{(n-1)^2}\left\{\sum_{i=2}^n m(u,t_{i-1})m(v,t_{i-1})r(t_i,u,v) + m(u,t_i)m(v,t_i)r(t_{i-1},u,v)\right.\\ &\quad \left. +\; r(t_i,u,v)r(t_{i-1},u,v) + \sum_{i=3}^n m(u,t_i)m(v,t_{i-2})r(t_{i-1},u,v) + m(v,t_i)m(u,t_{i-2})r(t_{i-1},u,v)\right\},\end{aligned}$$

and we therefore obtain $B_1 = \tilde{B}_1(u,v) + \tilde{B}_1(v,u)$ with

$$\begin{aligned}\tilde{B}_1(u,v) &= \frac{1}{(n-1)^2}\left\{\sum_{i=2}^n m(u,t_{i-1})^2 r(t_i,u,u) - m(u,t_{i-1})m(v,t_{i-1})r(t_i,u,v)\right.\\ &\quad + m(u,t_i)^2 r(t_{i-1},u,u) - m(u,t_i)m(v,t_i)r(t_{i-1},u,v) + r(t_i,u,u)r(t_{i-1},u,u) - r(t_i,u,v)r(t_{i-1},u,v)\\ &\quad \left. +\; 2\sum_{i=3}^n m(u,t_i)m(u,t_{i-2})r(t_{i-1},u,u) - m(u,t_i)m(v,t_{i-2})r(t_{i-1},u,v)\right\}.\end{aligned}$$

A typical summand in $\tilde{B}_1$ can be estimated by

$$|m(u,t_{i-1})^2 r(t_i,u,u) - m(u,t_{i-1})m(v,t_{i-1})r(t_i,u,v)| \le C|u-v|^\gamma$$

using the Lipschitz property of the functions r and m. All other summands are treated similarly, and we obtain

$$B_1 = B_1(u,v) \le \frac{1}{n-1} C|u-v|^\gamma,$$

which proves assertion (2.8) and completes the proof of Theorem 2.1. □

Remark 2.2. The assertion of Theorem 2.1 remains also valid, if the general hypothesis (1.3) of nonlinear functional regression models has to be tested, and we will indicate the arguments for proving this assertion here briefly. First note that the estimate $\hat{A}_{u,0}$ can be rewritten as

$$\hat{A}_{u,0} = \frac{1}{n}\sum_{i=1}^n Y_i^2(u) - \hat{\sigma}_u^2 + o_p\left(\frac{1}{\sqrt{n}}\right),$$

where

$$\hat{\sigma}_u^2 = \frac{1}{2(n-1)}\sum_{i=2}^n (Y_i(u) - Y_{i-1}(u))^2$$

denotes an estimate of the integrated variance

$$\int_0^1 \mathrm{Var}\ (\varepsilon(u,t))h(t)\mathrm{d}t = \int_0^1 r^2(t,u,u)h(t)\mathrm{d}t$$

at the point $u \in [0,1]$. Now a straightforward calculation shows that the estimate $\hat{M}_u^2$ is essentially the sum of squared residuals, i.e.

$$\hat{M}_u^2 = \min_\beta \frac{1}{n}\sum_{i=1}^n \left(Y_i(u) - \beta^T f(u,t_i)\right)^2 - \hat{\sigma}_u^2 + o_p\left(\frac{1}{\sqrt{n}}\right) \tag{2.10}$$

uniformly with respect to $u \in [0,1]$. Obviously, this concept can be easily generalized to the problem of testing the hypothesis of a nonlinear functional regression model. To be precise we assume that for each $u \in [0,1]$ the function $g_u : t \mapsto g(u,t,\beta_u)$ satisfies the standard regularity conditions of a nonlinear regression model (see for example [17] or [28]). In particular we assume the set $\Theta \subset \mathbb{R}^k$ is a compact set with non-empty interior and that for all $u, t \in [0,1]$ the function

$g(u,t,\beta)$ is twice continuously differentiable w.r.t. β

and satisfies

$g(u,\cdot,\beta), g(\cdot,t,\beta) \in \mathrm{Lip}_\gamma^{unif}[0,1]$.

We recall the definition (2.6) of the parameter corresponding to best L^2-approximation of the function $m(u,\cdot) : [0,1] \to \mathbb{R}$ by parametric functions of the form $\{g(u,\cdot,\beta_u) \mid \beta_u \in \Theta\}$, where we assume for each $u \in [0,1]$ the existence of the minimum $\beta_0(u)$ at a unique interior point of the compact space Θ. The L^2-distance between the function $m(u,\cdot)$ and its best approximation $g(u,\cdot,\beta_0(u))$ in the parametric class is now defined by

$$M_u^2 = \int_0^1 (m(u,t) - g(u,t,\beta_0(u)))^2\, h(t)\mathrm{d}t.$$

In order to investigate whether the hypothesis (1.3) is satisfied let for each $u \in [0,1]$

$$\hat{\beta}_0(u) = \mathrm{arginf}_\beta \sum_{i=1}^n (Y_i(u) - g(u,t_i,\beta))^2 \tag{2.11}$$

denote the nonlinear least squares estimate (here and throughout this paper it is assumed that the infimum in (2.11) is attained at a unique interior point of $\Theta \subset \mathbb{R}^k$) and observing (2.10) we obtain as the analogue of (2.3) the statistic

$$\hat{T}_u^2 = \frac{1}{n}\sum_{i=1}^n (Y_i(u) - g(u,t_i,\hat{\beta}_0(u)))^2 - \hat{\sigma}_u^2. \tag{2.12}$$

It follows by similar arguments as in [6] that

$$\hat{T}_u^2 = \frac{1}{n}\sum_{i=2}^n \varepsilon_i(u,t_i)\varepsilon_i(u,t_{i-1}) - \frac{2}{n}\sum_{i=1}^n (m(u,t_i) - g(u,t_i,\beta_0(u)))\varepsilon(u,t_i) + M_u^2 + o_p\left(\frac{1}{\sqrt{n}}\right)$$

uniformly with respect to $u \in [0,1]$, and a similar reasoning as presented in the proof of Theorem 2.1 shows that

$$\sqrt{n}(\hat{T}_u^2 - M_u^2) \Longrightarrow G,$$

where the covariance structure of the Gaussian process G is specified in (2.5). The details are omitted for the sake of brevity.

Note that the null hypothesis (1.3) is satisfied if and only if $M_u^2 = 0$ for all $u \in [0,1]$. Consequently, a consistent test can be obtained by rejecting the null hypotheses for large values of a Cramér-von-Mises or a Kolmogoroff–Smirnov functional of the process $\{\hat{T}_u\}_{u\in[0,1]}$. Under the null hypothesis the covariance kernel of the limiting process G in Theorem 2.1 reduces to

$$k(u,v) \stackrel{H_0}{=} \int_0^1 r^2(t,u,v)h(t)\mathrm{d}t \tag{2.13}$$

and by the continuous mapping theorem it follows that the statistic

$$\sqrt{n}\int_0^1 \hat{M}_u^2 \mathrm{d}u$$

converges weakly to a centered normal distribution with variance $\int_0^1\int_0^1 k(u,v)\mathrm{d}u\mathrm{d}v$. Therefore, it remains to estimate the asymptotic variance, and we propose to use

$$\hat{s}_n^2 = \int_0^1\int_0^1 \hat{k}(u,v)\mathrm{d}u\mathrm{d}v,$$

where the estimate of the covariance kernel $k(u, v)$ is defined by

$$\hat{k}(u, v) = \frac{1}{4(n-3)} \sum_{i=2}^{n-2} S_i(u)S_i(v)S_{i+2}(u)S_{i+2}(v). \tag{2.14}$$

with $S_i(u) = Y_i(u) - Y_{i-1}(u)$. The following result shows that under the null hypothesis the statistic $\hat{s}_n^2$ is a consistent estimate of the asymptotic variance. The technical details of the proof are given in Appendix.

Proposition 2.1. *Under the assumptions of Theorem 2.1 we have*

$$\hat{k}(u, v) = \int_0^1 r^2(t, u, v)h(t)\mathrm{d}t + O_p(n^{-1/2})$$

uniformly with respect to $u, v \in [0, 1]$.

Theorem 2.1, Proposition 2.1 and (2.13) provide an asymptotic level α test by rejecting the null hypothesis (1.3) if

$$M_n = \frac{\sqrt{n}}{\sqrt{\int_0^1 \int_0^1 \hat{k}(u, v)\,\mathrm{d}u\,\mathrm{d}v}} \int_0^1 \hat{M}_u^2\,\mathrm{d}u > u_{1-\alpha}, \tag{2.15}$$

where $u_{1-\alpha}$ denotes the $(1-\alpha)$ quantile of the standard normal distribution. Due to Theorem 2.1 this test is consistent under fixed alternatives satisfying $\int_0^1 M_u^2\,\mathrm{d}u > 0$. Furthermore, it can detect local alternatives converging to the null hypothesis at a rate $n^{-1/4}$. More precisely, suppose that

$$Y_i(u) = \beta_0^T(u)f(u, t_i) + \frac{1}{n^{1/4}}c(u, t_i) + \varepsilon(u, t_i)$$

with some fixed function $c : [0, 1]^2 \to \mathbb{R}$. We assume that for fixed u or fixed t the function c is uniformly Lipschitz continuous in the other argument. In this case, mimicking the arguments from the proof of Theorem 2.1, it can be shown that

$$\left(\sqrt{n}\hat{M}_u^2\right)_{u\in[0,1]} \Longrightarrow \left(G_u + \int_0^1 (c(u, t) - (P_u c)(t))^2\, h(t)\,\mathrm{d}t\right)_{u\in[0,1]}$$

in $C[0, 1]$, where $P_u c$ denotes the orthogonal projection of $c(u, \cdot)$ on $\mathrm{span}\{f_1(u, \cdot), \ldots, f_k(u, \cdot)\}$. The finite sample properties of the test and a corresponding bootstrap version will be illustrated in Section 5.

Remark 2.3. It is worthwhile to mention that the results can be extended to the case of dependent data, where the errors $\varepsilon(u, t_i)$ are generated by a stationary causal process

$$\varepsilon(u, t_i) = \sum_{j=0}^{\infty} b_j e_{i-j}(u),$$

where $\{e_j(u) \mid u \in [0, 1]\}_{j\in\mathbb{N}_0}$ is a sequence of independent identically distributed stochastic processes with zero mean, such that the autocovariance function $\gamma_k(u) = E[\varepsilon(u, t_1)\varepsilon(u, t_{k+1})]$ is absolutely summable and additionally for all $u \in [0, 1]$ the condition

$$\sum_{s=-\infty}^{\infty} |s||\gamma_s(u)| < \infty$$

is satisfied. For the sake of brevity we do not present details here but refer to the work of González Manteiga and Vilar Fernández [18] and Dette and Biedermann [11] who considered this dependence structure in the context of testing for a parametric form of the regression function on the basis of kernel methods.

3. A test for the hypothesis (1.4)

We now consider the problem of testing the hypothesis (1.4) in the functional regression model defined by (1.1) and assume that n independent observations are available. For this purpose we define for fixed $t \in [0, 1]$ the L^2-distance

$$M_t^2 = \inf_{\gamma_t} \int_0^1 (m(u, t) - h(u, t, \gamma_t))^2\,\mathrm{d}u. \tag{3.1}$$

We only deal with the linear case, that is

$$h(u, t, \gamma(t)) = \gamma(t)^T f(u, t)$$

for some given regression functions $f(u,t) = (f_1(u,t), \dots, f_k(u,t))$ and denote by $\gamma_0(t)$ the function, which yields to the minimal values in (3.1). As a global measure of deviance from the null hypothesis we consider the functional

$$M^2 = \int_0^1 M_t^2\, h(t)\, dt, \tag{3.2}$$

and obviously the hypothesis $H_0 : M^2 = 0$ is equivalent to (1.4).

Similarly as in Section 2, standard Hilbert space theory shows that the distance M_t^2 can be expressed as a ratio of two Gramian determinants

$$M_t^2 = \frac{\Gamma_t(m, f_1, \dots, f_k)}{\Gamma(f_1, \dots, f_k)}, \tag{3.3}$$

where $\Gamma_t(p_1, \dots, p_l) = \det((\langle p_i, p_j\rangle_t)_{i,j=1}^l)$ and the inner products are now calculated with respect to the variable u, that is

$$\langle f, g\rangle_t = \int_0^1 f(u,t) g(u,t)\, du.$$

For the time t_i we can "estimate" the entries of the matrix in the numerator of (3.3) by

$$\hat{B}_{i,0} = \int Y_i(u) Y_{i-1}(u)\, du,$$
$$\hat{B}_{i,p} = \int Y_i(u) f_p(u, t_i)\, du,$$
$$\hat{C}_{i,p} = \int Y_{i-1}(u) f_p(u, t_{i-1})\, du = \hat{B}_{i-1,p},$$

and define

$$\hat{M}_{t_i}^2 = \frac{\begin{vmatrix} \hat{B}_{i,0} & \hat{B}_{i,1} & \cdots & \hat{B}_{i,k} \\ \hat{C}_{i,1} & \langle f_1, f_1\rangle_{t_i} & \cdots & \langle f_1, f_k\rangle_{t_i} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{C}_{i,k} & \langle f_k, f_1\rangle_{t_i} & \cdots & \langle f_k, f_k\rangle_{t_i} \end{vmatrix}}{\begin{vmatrix} \langle f_1, f_1\rangle_{t_i} & \cdots & \langle f_1, f_k\rangle_{t_i} \\ \vdots & \ddots & \vdots \\ \langle f_k, f_1\rangle_{t_i} & \cdots & \langle f_k, f_k\rangle_{t_i} \end{vmatrix}} \tag{3.4}$$

as an estimator for $M_{t_i}^2$. Note that we estimate the entries in the upper first column by $\hat{C}_{i,p}$ rather than $\hat{B}_{i,p}$ in order to assure that the statistic $\hat{M}_{t_i}^2$ is asymptotically unbiased. However, because only one observation is made at time t_i, the variance of $\hat{M}_{t_i}^2$ is not converging to 0 with increasing sample size. As a consequence, the statistic $\hat{M}_{t_i}^2$ is not a consistent estimate for $M_{t_i}^2$. Nevertheless, a consistent estimate for the measure defined in (3.2) can be obtained by averaging the quantities $\hat{M}_{t_i}^2$, that is

$$\hat{M}^2 = \frac{1}{n-1} \sum_{i=2}^n M_{t_i}^2. \tag{3.5}$$

Similarly, consistent estimates of M_t^2 at a particular point t can be obtained by local averages.

Theorem 3.1. *Under the assumptions of Section 2 the estimate $\hat{M}^2$ defined in (3.5) is consistent for $M^2 = \int_0^1 M_t^2\, dt$. More precisely, we have as $n \to \infty$*

$$\sqrt{n-1}(\hat{M}^2 - M^2) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2),$$

where the asymptotic variance is given by

$$\sigma^2 = \int_0^1 \bigg(\int_{[0,1]^2} \big(r(t,u,v) - (P_{u,t} r)(v)\big) \big(r(t,u,v) - (P_{v,t} r)(u)\big)\, du\, dv$$
$$+ 4 \int_{[0,1]^2} r(t,u,v)\, (m(u,t) - (P_t m)(u))\, (m(v,t) - (P_t m)(v))\, du\, dv \bigg) h(t) dt,$$

and $(P_t m)(u) = \gamma_{t,0}^T f(u, t)$ *and* $(P_{u,t} r)(v) = \gamma_{u,t,0}^T f(v, t)$ *denote the orthogonal projections of the function* $m(\cdot, t)$ *and* $r(u, \cdot, t)$ *on the set span* $\{f_1(\cdot, t), \dots, f_k(\cdot, t)\}$*, respectively, that is*

$$\int_0^1 (m(u,t) - \gamma_{t,0}^T f(u,t))^2 \, du = \inf_{\gamma_t} \int_0^1 (m(u,t) - \gamma_t^T f(u,t))^2 \, du = M_t^2,$$
$$\int_0^1 (r(t,u,v) - \gamma_{u,t,0}^T f(v,t))^2 \, dv = \inf_{\gamma_{u,t}} \int_0^1 (r(t,u,v) - \gamma_{u,t}^T f(v,t))^2 \, dv.$$

Proof of Theorem 3.1. Without loss of generality we may assume that the functions $f_1, \dots, f_k$ are orthonormal and therefore the minimal distance in (3.3) and its estimator defined in (3.4) simplify to

$$M_{t_i}^2 = \langle m, m\rangle_{t_i} - \sum_{p=1}^k \langle m, f_p\rangle_{t_i}^2,$$
$$\hat{M}_{t_i}^2 = \hat{B}_{i,0} - \sum_{p=1}^k \hat{B}_{i,p}\hat{C}_{i,p},$$

respectively. A careful calculation of the moments of the random variables in the latter expression yields

$$E[\hat{B}_{i,0}] = \langle m, m\rangle_{t_i} + O(n^{-\gamma}),$$
$$E[\hat{B}_{i,p}\hat{C}_{i,p}] = \langle m, f_p\rangle_{t_i}^2 + O(n^{-\gamma}),$$
$$\mathrm{Var}(\hat{B}_{i,0}) = \int r(t_i,u,v)^2 \, du\, dv + 2\int r(t_i,u,v) m(u,t_i) m(v,t_i)\, du\, dv + O(n^{-\gamma}),$$
$$\begin{aligned}\mathrm{Cov}(\hat{B}_{i,p}\hat{C}_{i,p}, \hat{B}_{i,q}\hat{C}_{i,q}) &= \int r(t_i,u,v) f_p(u,t_i) f_q(v,t_i)\, du\, dv \Big(2\langle m, f_p\rangle_{t_i} \langle m, f_q\rangle_{t_i} \\ &\quad + \int r(t_i,u,v) f_p(u,t_i) f_q(v,t_i)\, du\, dv \Big) + O(n^{-\gamma}),\end{aligned}$$
$$\begin{aligned}\mathrm{Cov}(\hat{B}_{i,0}, \hat{B}_{i,p}\hat{C}_{i,p}) &= 2\int r(t_i,u,v) m(u,t_i) f_p(v,t)\, du\, dt \ \langle m, f_p\rangle_{t_i} \\ &\quad + \int r(t_i,u,v) r(t_i,u,w) f_p(v,t_i) f_p(w,t_i)\, du\, dv\, dw + O(n^{-\gamma}),\end{aligned}$$
$$\mathrm{Cov}(\hat{B}_{i,0}, \hat{B}_{i-1,0}) = \int r(t_i,u,v) m(u,t_i) m(v,t_i)\, du\, dv + O(n^{-\gamma}),$$
$$\begin{aligned}\mathrm{Cov}(\hat{B}_{i,0,}, \hat{B}_{i-1,p}\hat{C}_{i-1,p}) &= \int r(t_i,u,v) m(u,t_i) f_p(v,t_i)\, du\, dv \ \langle m, f_p\rangle_{t_i} + O(n^{-\gamma}) \\ &= \mathrm{Cov}(\hat{B}_{i-1,0}, \hat{B}_{i,p}\hat{C}_{i,p}),\end{aligned}$$
$$\mathrm{Cov}(\hat{B}_{i,p}\hat{C}_{i,p}, \hat{B}_{i-1,q}\hat{C}_{i-1,q}) = \int r(t_i,u,v) f_p(u,t_i) f_q(v,t_i)\, du\, dv \ \langle m, f_p\rangle_{t_i} \langle m, f_q\rangle_{t_i} + O(n^{-\gamma}).$$

The sequence $\hat{M}_{t_2}^2, \dots, \hat{M}_{t_n}^2$ forms a triangular array of one-dependent random variable and as a consequence all covariances corresponding to a lag larger than one vanish. Therefore the variance of the standardized mean

$$\sigma_n^2 = \mathrm{Var}\left(\frac{1}{\sqrt{n-1}} \sum_{i=2}^n \hat{M}_{t_i}^2 \right)$$

is given by

$$\begin{aligned}\sigma_n^2 &= \frac{1}{n-1} \sum_{i=2}^n \Bigg\{ \mathrm{Var}(B_{i,0}) + \sum_{p,q=1}^k \mathrm{Cov}(\hat{B}_{i,p}\hat{C}_{i,p}, \hat{B}_{i,q}\hat{C}_{i,q}) - 2\sum_{p=1}^k \mathrm{Cov}(\hat{B}_{i,0}, \hat{B}_{i,p}\hat{C}_{i,p}) \\ &\quad + 2\,\mathrm{Cov}(\hat{B}_{i,0}, B_{i-1,0}) - 2\sum_{p=1}^k \mathrm{Cov}(\hat{B}_{i,0}, \hat{B}_{i-1,p}\hat{C}_{i-1,p}) \\ &\quad - 2\sum_{p=1}^k \mathrm{Cov}(\hat{B}_{i-1,0}, \hat{B}_{i,p}\hat{C}_{i,p}) + 2\sum_{p,q=1}^k \mathrm{Cov}(\hat{B}_{i,p}\hat{C}_{i,p}, \hat{B}_{i-1,q}\hat{C}_{i-1,q}) \Bigg\} + O(n^{-\gamma})\end{aligned}$$

$$
\begin{aligned}
&= \int_0^1 \Bigg\{ \int r(t,u,v)^2 \,\mathrm{d}u\,\mathrm{d}v - 2\sum_{p=1}^{k} \int r(t,u,v) r(t,u,w) f_p(v,t) f_q(w,t)\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w \\
&\quad + \sum_{p,q=1}^{k} \left(\int r(t,u,v) f_p(u,t) f_q(v,t)\,\mathrm{d}u\,\mathrm{d}v \right)^2 \\
&\quad + 4\int r(t,u,v) m(u,t) m(v,t)\,\mathrm{d}u\,\mathrm{d}v - 8\sum_{i=1}^{p} \int r(t,u,v) m(u,t) f_p(v,t)\,\mathrm{d}u\,\mathrm{d}v\,\langle m, f_p\rangle_t \\
&\quad + 4\sum_{p,q=1}^{k} \int r(t,u,v) f_p(u,t) f_q(v,t)\,\mathrm{d}u\,\mathrm{d}v\,\langle m, f_p\rangle_t\,\langle m, f_q\rangle_t \Bigg\}\,\mathrm{d}t + O(n^{-\gamma}) \\
&= \sigma^2 + O(n^{-\gamma}).
\end{aligned}
$$

Here the last equality uses the fact that under the assumption of orthonormality the orthogonal projection $P_t m$ and $P_{u,t} r$ are given by

$$
\begin{aligned}
(P_t m)(u) &= \sum_{p=1}^{k} \langle m, f_p\rangle_t f_p(u,t), \\
(P_{u,t} r)(v) &= \sum_{p=1}^{k} \langle r(\cdot, u, \cdot), f_p\rangle_t f_p(v,t).
\end{aligned}
$$

The assertion of the theorem now follows by the classical central limit theorem for m-dependent random variables (see [24]). $\square$

Under the null hypothesis the variance of the limiting normal distribution simplifies to

$$
\sigma^2 \overset{H_0}{=} \int_0^1 \left(\int_{[0,1]^2} \big(r(u,v,t) - (P_{u,t} r)(v)\big)\big(r(u,v,t) - (P_{v,t} r)(u)\big)\,\mathrm{d}(u,v) \right) h(t)\,\mathrm{d}t.
$$

We propose to estimate this variance by

$$
\begin{aligned}
\hat{\sigma}^2 &= \frac{1}{4(n-3)} \sum_{i=2}^{n-2} \int_{[0,1]^2} S_i(u) \left\{ S_i(v) - \int_0^1 S_i(x) f(x,t_i)^T \,\mathrm{d}x A_i^{-1} f(v,t_i) \right\} \\
&\quad \times S_{i+2}(v) \left\{ S_{i+2}(u) - \int_0^1 S_{i+2}(x) f(x,t_{i+2})^T \,\mathrm{d}x\, A_{i+2}^{-1} f(u,t_{i+2}) \right\} \mathrm{d}(u,v),
\end{aligned}
$$

where $S_i(u) = Y_i(u) - Y_{i-1}(u)$ and $A_i = \int_0^1 f(u,t_i) f(u,t_i)^T \,\mathrm{d}u \in \mathbb{R}^{k\times k}$. Observing that the orthogonal projection $(P_{u,t} r)(v)$ is given by

$$
P_{u,t} r(v) = \gamma_{u,t,0}^T f(v,t) = \int_0^1 r(u,x,t) f(x,t)^T \,\mathrm{d}x \left(\int_0^1 f(x,t) f(x,t)^T \,\mathrm{d}x \right)^{-1} f(v,t)
$$

it follows by a similar but rather tedious calculation as in the proof of Theorem 3.1 that $\hat{\sigma}^2$ is a $\sqrt{n}$-consistent estimator for σ^2. For example it can be shown that

$$
E[S_i(u) S_i(v)] = 2r(u,v,t_i) + O(n^{-\gamma}),
$$

and the same arguments for the other terms yield that $\hat{\sigma}^2$ is asymptotically unbiased with a rate $o(n^{-1/2})$. Similarly the variance of $\hat{\sigma}^2$ can be shown to be of order n^{-1}. Therefore we obtain an asymptotic level α test for the hypothesis (1.4) by rejecting H_0 if

$$
\sqrt{\frac{n-1}{\hat{\sigma}^2}}\hat{M}^2 > u_{1-\alpha}, \tag{3.6}
$$

where $u_{1-\alpha}$ denotes the $(1-\alpha)$ quantile of the standard normal distribution. Due to Theorem 3.1 this test is consistent against all alternatives satisfying $M^2 > 0$. Furthermore, it can detect local alternatives converging to the null hypothesis at a rate $n^{-1/4}$. More precisely, suppose that

$$
Y_i(u) = \gamma_0^T(u) f(u,t_i) + \frac{1}{n^{1/4}} c(u,t_i) + \varepsilon(u,t_i)
$$

with some fixed function $c : [0,1]^2 \to \mathbb{R}$. We assume that for fixed u or fixed v the function c is uniformly Lipschitz continuous in the other argument. In this case, mimicking the arguments from the proof of Theorem 3.1, it can be shown that

$$\sqrt{n}\,\hat{M}^2 \xrightarrow{\mathcal{D}} \mathcal{N}(\mu, \sigma^2),$$

where $\mu = \int_0^1 \int_0^1 (c(u,t) - (P_t c)(u))^2 \,\mathrm{d}u\, h(t)\mathrm{d}t$.

4. Testing homoscedasticity

In this section we address the problem of testing the hypothesis (1.5) of homoscedastic errors in the functional regression model (1.1). Motivated by the discussion in Sections 2 and 3 we propose the following measure of heteroscedasticity at a point $(u, v) \in [0, 1]^2$

$$\tau^2(u,v) = \min_{a\in\mathbb{R}} \|r(\cdot,u,v) - a\|_{u,v}^2 = \int_0^1 r^2(t,u,v)h(t)\,\mathrm{d}t - \left(\int_0^1 r(t,u,v)h(t)\,\mathrm{d}t\right)^2, \tag{4.1}$$

where $\|f(\cdot,u,v)\|_{u,v}^2 = \int f^2(t,u,v)\,\mathrm{d}t$. Note that $\tau^2(u,v) = 0$ a.e. if and only if the covariance function does not depend on t, that is the hypothesis (1.5) of homoscedasticity is valid. An estimator for the quantity $\int_0^1 r^2(t,u,v)h(t)\,\mathrm{d}t$ in (4.1) has been proposed in (2.14), and for the second term we will use a similar estimate based on the statistic

$$\tilde{k}(u,v) = \frac{1}{2(n-1)} \sum_{i=2}^{n} S_i(u)S_i(v),$$

where $S_i(u) = Y_i(u) - Y_{i-1}(u)$. We therefore obtain as an estimator of the process $\{\tau^2(u,v)\}_{u,v\in[0,1]}$

$$\hat{\tau}_n^2(u,v) = \frac{1}{4(n-3)} \sum_{i=2}^{n-2} S_i(u)S_i(v)S_{i+1}(u)S_{i+2}(v) - \left(\frac{1}{2(n-1)} \sum_{i=2}^{n} S_i(u)S_i(v)\right)^2.$$

The asymptotic properties of this random variable are specified in the following result.

Theorem 4.1. *Assume that the third and fourth moments*

$$d_1(t,u,v,w) = E[\varepsilon(u,t)\varepsilon(v,t)\varepsilon(w,t)]$$
$$d_2(t,u,v,w,x) = E[\varepsilon(u,t)\varepsilon(v,t)\varepsilon(w,t)\varepsilon(x,t)]$$

of the error process $\varepsilon(u,t)$ exist and are elements of $\mathrm{Lip}_\gamma^{unif}[0,1]$ for every argument. If the assumptions of Section 2 are satisfied we have as $n \to \infty$

$$4\sqrt{n}(\hat{\tau}_n^2(u,v) - \tau^2(u,v)) \Longrightarrow G$$

in $C[0,1]^2$. Here G is a centered Gaussian field on $[0,1]^2$ whose covariance structure under the null hypothesis of homoscedasticity is given by

$$\begin{aligned}
k&((u_1,v_1),(u_2,v_2)) := \mathrm{Cov}\,(G(u_1,v_1), G(u_2,v_2)) \\
&= 6D_2^{(2)}(u_1,v_1,u_2,v_2) - 12D_2^{(r,1,1)}(u_1,v_1,u_2,v_2) + 8D_2^{(r,1,1)}(u_1,u_2,v_1,v_2) + 8D_2^{(r,1,1)}(u_1,v_2,v_1,u_2) \\
&\quad + 6J(u_1,v_1,u_2,v_2,u_1,v_1,u_2,v_2) + 4J(u_1,u_2,v_1,v_2,u_1,u_2,v_1,v_2) \\
&\quad + 4J(u_1,v_2,v_1,u_2,u_1,v_2,v_1,u_2) - 8J(u_1,v_1,u_2,v_2,u_1,u_2,v_1,v_2) \\
&\quad - 8J(u_1,v_2,u_2,v_2,u_1,v_2,v_1,u_2) + 8J(u_1,v_1,u_2,v_2,u_1,v_2,v_1,u_2) \\
&\quad + 2D_1^{(r)}(u_1,u_2,v_1,v_2) + 2D_1^{(r)}(u_1,v_2,v_1,u_2) + 2D_1^{(r)}(v_1,u_2,u_1,v_2) + 2D_1^{(r)}(v_1,v_2,u_1,u_2)
\end{aligned}$$

where the following notations have been used

$$D_2^{(2)}(u_1,v_1,u_2,v_2) = \int_0^1 \mathrm{d}_2(t,u_1,v_1,u_2,v_2)^2 h(t)\,\mathrm{d}t$$

$$D_2^{(r,i,j)}(u_1,v_1,u_2,v_2) = r(u_1,v_1)^i r(u_2,v_2)^j \int_0^1 \mathrm{d}_2(t,u_1,v_1,u_2,v_2)h(t)\,\mathrm{d}t$$

$$J(u_1,v_1,u_2,v_2,u_3,v_3,u_4,v_4) = \prod_{i=1}^{4} r(u_i,v_i)$$

$$D_1^{(r)}(u_1,v_1,u_2,v_2) = r(u_1,v_1)\int_0^1 \mathrm{d}_1(t,v_1,u_2,v_2)\mathrm{d}_1(t,u_1,u_2,v_2)h(t)\,\mathrm{d}t.$$

Proof of Theorem 4.1. The proof follows along similar lines as the proof of Theorem 2.1, establishing weak convergence of finite dimensional distributions and tightness of the sequence

$$\{4\sqrt{n}(\hat{\tau}_n^2(u,v)-\tau^2(u,v))\}_{u,v\in[0,1]}.$$

For this reason only the main steps are indicated in the subsequent discussion. A careful inspection of the results in the proof of Lemmas 6.2 and 6.3 in [12] yields to the following decomposition into a sum of 4-dependent random variables and a stochastic remainder of order $n^{-\frac{1}{2}}$

$$\hat{\tau}_n^2(u,v)-\tau^2(u,v)=\frac{1}{4(n-3)}\sum_{j=2}^{n-2}W_j(u,v)+o_P(n^{-\frac{1}{2}})$$

(uniformly with respect to (u,v)), where

$$\begin{aligned}
W_j(u,v)&=Z_j(u,v)\left\{Z_{j+2}(u,v)+4\delta_j(u,v)\right\},\\
Z_j(u,v)&=\Delta\varepsilon_{u,j-1,j}\Delta\varepsilon_{v,j-1,j}-E_j(u,v),\\
E_j(u,v)&=E[\Delta\varepsilon_{u,j-1,j}\Delta\varepsilon_{v,j-1,j}]=2r(t_j,u,v)+O(n^{-\gamma}),\\
\delta_j(u,v)&=r(t_j,u,v)-\frac{1}{n}\sum_{i=1}^{n}r(t_i,u,v).
\end{aligned}$$

A straightforward but tedious calculation shows that the covariance structure of the random variables $W_j(u,v)$ is given by

$$\begin{aligned}
&\mathrm{Cov}(W_j(u_1,v_1),W_j(u_2,v_2))\\
&\quad=4(d_2(t_j,u_1,v_1,u_2,v_2)+r(t_j,u_1,v_1)r(t_j,u_2,v_2)+r(t_j,u_1,u_2)r(t_j,v_1,v_2)+r(t_j,u_1,v_2)r(t_j,v_1,u_2))^2\\
&\qquad+16(d_2(t_j,u_1,v_1,u_2,v_2)+r(t_j,u_1,v_1)r(t_j,u_2,v_2)+r(t_j,u_1,u_2)r(t_j,v_1,v_2)\\
&\qquad+r(t_j,u_1,v_2)r(t_j,v_1,u_2))(2\delta_j(u_1,v_1)\delta_j(u_2,v_2))-r(t_j,u_1,v_1)r(t_j(u_2,v_2))\\
&\qquad+16r^2(t_j,u_1,v_1)r^2(t_j,u_2,v_2)-64r(t_j,u_1,v_1)r(t_j,u_2,v_2)\delta_j(u_1,v_1)\delta_j(u_2,v_2)+O(n^{-\gamma}),\\
&\mathrm{Cov}(W_j(u_1,v_1),W_{j+1}(u_2,v_2))\\
&\quad=(d_2(t_j,u_1,v_1,u_2,v_2))^2-2d_2(t_j,u_1,v_1,u_2,v_2)r(t_j,u_1,v_1)r(t_j,u_2,v_2)+r^2(t_j,u_1,v_1)r^2(t_j,u_2,v_2)\\
&\qquad+d_1(t_j,v_1,u_2,v_2)d_1(t_j,u_1,v_1,u_2)r(t_j,u_1,u_2)+d_1(t_j,v_1,u_2,v_2)d_1(t_j,u_1,v_1,u_2)r(t_j,u_1,v_2)\\
&\qquad+d_1(t_j,u_1,u_2,v_2)d_1(t_j,u_1,v_1,v_2)r(t_j,v_1,u_2)+d_1(t_j,u_1,u_2,v_2)d_1(t_j,u_1,v_1,u_2)r(t_j,v_1,v_2)\\
&\qquad-8\delta_j(u_2,v_2)d_1(t_j,u_1,v_1,v_2)d_1(t_j,u_1,v_1,u_2)+16\delta_j(u_1,v_1)\delta_j(u_2,v_2)\\
&\qquad\times\left(d_2(t_j,u_1,v_1,u_2,v_2)-r(t_j,u_1,v_1)r(t_j,u_2,v_2)\right)+O(n^{-\gamma})
\end{aligned}$$

and

$$\mathrm{Cov}(W_j(u_1,v_1),W_i(u_2,v_2))=0\quad\text{for } |i-j|\geq 2.$$

The dominating sum

$$A_n(u,v)=\frac{1}{4(n-3)}\sum_{j=2}^{n-2}W_j(u,v)$$

therefore has asymptotic covariance

$$\begin{aligned}
16n\,\mathrm{Cov}(A_n(u_1,v_1),A_n(u_2,v_2))&=\frac{1}{n}\sum_{j=2}^{n-2}\mathrm{Cov}(W_j(u_1,v_1),W_j(u_2,v_2))+\mathrm{Cov}(W_j(u_1,v_1),W_{j+1}(u_2,v_2))\\
&\quad+\mathrm{Cov}(W_j(u_2,v_2),W_{j+1}(u_1,v_1))+o(1)\\
&=k\left((u_1,v_1),(u_2,v_2)\right)+o(1).
\end{aligned}$$

The last equality is obtained using the Lipschitz continuity of the regression functions. Finally the validation of tightness follows along similar lines as in the proof of Theorem 2.1 by a tedious calculation of a corresponding moment condition for Gaussian fields (see e.g. [3]) and is therefore omitted. □

5. Finite sample properties

In this section we study the finite sample properties of the tests proposed in the previous sections. Our first example considers the linear hypothesis

$$H_0: m(u,t)=g(u,t,\beta(u))=\beta(u)f(u,t),$$

Table 1
Simulated rejection probabilities of the test (2.15) under the null hypothesis $H_0 : m(u, t) = f_i(u, t), i = 1, 2$, where the regression functions f_1 and f_2 are given in (5.1) and (5.2), respectively.

	n	Mean	Var	0.15	0.1	0.05	0.01
	25	−0.2484	1.4008	0.1192	0.0920	0.0664	0.0334
	50	−0.1359	1.2099	0.1336	0.1022	0.0632	0.0262
$f_1(u, t)$	100	−0.0975	1.0773	0.1328	0.0954	0.0544	0.0202
	200	−0.0290	1.0516	0.1464	0.1062	0.0674	0.0208
	500	−0.0373	1.0537	0.1514	0.1064	0.0578	0.0170
	25	−1.28	1.4253	0.0382	0.0264	0.0152	0.0056
	50	−0.6477	1.2543	0.0726	0.0538	0.0372	0.0146
$f_2(u, t)$	100	−0.3862	1.1260	0.0886	0.0676	0.0434	0.0164
	200	−0.2797	1.0379	0.1014	0.0718	0.0402	0.0102
	500	−0.1455	1.0267	0.1226	0.0802	0.0444	0.0134

Table 2
Simulated rejection probabilities of the test (2.15) under the alternative $H_1 : m(u, t) = f_i(u, t) + 1/2 \exp(t), i = 1, 2$, where the regression functions f_1 and f_2 are given in (5.1) and (5.2), respectively.

	n	Mean	Var	0.15	0.1	0.05	0.01
	25	2.52	4.32	0.764	0.705	0.615	0.480
	50	3.72	3.92	0.936	0.914	0.874	0.740
$f_2(u, t)$	100	5.13	3.68	0.997	0.995	0.990	0.955
	200	7.18	3.47	1	1	1	1
	500	11.39	3.60	1	1	1	1
	25	7.99	23.49	0.981	0.977	0.970	0.939
	50	13.31	25.02	1	1	1	1
$f_2(u, t)$	100	19.49	26.67	1	1	1	1
	200	27.52	27.33	1	1	1	1
	500	44.45	29.01	1	1	1	1

where $f : [0, 1]^2 \to \mathbb{R}$ is some given function and $\beta : [0, 1] \to \mathbb{R}$ (i.e. $k = 1$). The discussion following the proof of Theorem 2.1 states that under the null hypothesis H_0, the statistic M_n defined in (2.15) converges weakly to a standard normal distribution. We reject the hypothesis H_0 if the inequality (2.15) is satisfied. In order to study the approximation of the nominal level and the power of this asymptotic level α test 5000 replications with different functions f have been performed. The error terms $\varepsilon(u, t_i)$ are assumed to be i.i.d. Brownian Motions, i.e. $r(t, u, v) = u \wedge v$, which implies that the model is homoscedastic and the parameter function β is chosen as $\beta \equiv 1$. The results under the null hypothesis are presented in Table 1 for the functions

$$f_1(u, t) = (-1 + 2u) + 2(1 - u)t \tag{5.1}$$

$$f_2(u, t) = (1 + u) \cos(2\pi t). \tag{5.2}$$

It can be seen that the nominal level of the test is well approximated in most cases. For the function $f_1(u, t) = (-1 + 2u) + (2 - 2u)t$ the approximation is very accurate for sample sizes larger than $n = 100$, for smaller values the level is either overestimated (if the nominal level is smaller than $\alpha = 0.1$) or underestimated (if the nominal level is larger than $\alpha = 0.1$). In the case where we use the function $f_2(u, t) = (1 + u) \cos(2\pi t)$ we underestimate the level, with the tendency to get better approximations for larger sample sizes.

For the investigation of the power of the test we consider the functions f_i defined in (5.1) and (5.2) with two additive alternatives, that is

$$m(u, t) = f_i(u, t) + \frac{1}{2} \exp(t) \tag{5.3}$$

$$m(u, t) = f_i(u, t) + \sin(2\pi t) \tag{5.4}$$

with $i = 1, 2$. The corresponding results are presented in Tables 2 and 3. We observe reasonable rejection probabilities for all sample sizes and both choices of f_i.

Note that for sample sizes $n = 25$ and $n = 50$ the approximation of the nominal level is less accurate. In these cases we propose a wild bootstrap procedure to obtain a more accurate test procedure [see [30]]. For this purpose we denote by $\hat{\beta}(u)$ the (point-wise) ordinary least square estimator of the function $\beta(u)$ and calculate the parametric residuals by

$$\hat{\varepsilon}(u, t_i) = Y_i(u) - \hat{\beta}(u) f_1(u, t_i) \tag{5.5}$$

for $i = 1, \dots, n$ and $u \in [0, 1]$. For $b = 1, \dots B$ with $B \in \mathbb{N}$ let v_i^{b*} be independent samples of a random variable V with a Laplacian distribution on the set $\{-1, 1\}$, and define the bootstrap sample as

$$Y_i^{b*}(u) = \hat{\beta}(u) f_1(u, t_i) + \varepsilon_i^{b*}(u); \quad i = 1, \dots, n, \tag{5.6}$$

Table 3
Simulated rejection probabilities of the test (2.15) under the alternative $H_1 : m(u,t) = f_i(u,t) + \sin(2\pi t)$, $i = 1, 2$, where the regression functions f_1 and f_2 are given in (5.1) and (5.2), respectively.

	n	Mean	Var	0.15	0.1	0.05	0.01
$f_1(u,t)$	25	11.48	33.036	1	1	0.9996	0.9984
	50	15.47	27.2397	1	1	1	1
	100	21.09	24.1177	1	1	1	1
	200	29.52	23.459	1	1	1	1
	500	45.92	22.6557	1	1	1	1
$f_2(u,t)$	25	5.047	11.9046	0.9398	0.9174	0.0876	0.7932
	50	8.645	14.3088	0.9988	0.9982	0.9966	0.9866
	100	12.51	14.4536	1	1	1	1
	200	17.69	13.7865	1	1	1	1
	500	27.69	13.3727	1	1	1	1

Table 4
Simulated rejection probabilities of the bootstrap test under the null hypothesis $H_0 : m(u,t) = f_i(u,t)$, $i = 1, 2$, where the regression functions f_1 and f_2 are given in (5.1) and (5.2), respectively.

	n	0.15	0.1	0.05	0.01
$f_1(u,t)$	25	0.15	0.108	0.055	0.020
	50	0.15	0.101	0.058	0.016
$f_2(u,t)$	25	0.158	0.108	0.057	0.020
	50	0.154	0.095	0.051	0.013

Table 5
Simulated rejection probabilities of the bootstrap test under the alternative $H_1 : m(u,t) = f_i(u,t) + 1/2\exp(t)$, $i = 1, 2$, where the regression functions f_1 and f_2 are given in (5.1) and (5.2), respectively.

	n	0.15	0.1	0.05	0.01
$f_1(u,t)$	25	0.856	0.792	0.640	0.440
	50	0.956	0.942	0.908	0.756
$f_2(u,t)$	25	0.996	0.990	0.980	0.926
	50	1	1	1	1

where

$$\varepsilon_i^{b*}(u) = v_i^{b*}\,\hat{\varepsilon}(u, t_i). \tag{5.7}$$

Note that we use a wild bootstrap with parametric residuals in order to avoid the choice of a smoothing parameter. For each $b \in \{1, \ldots, B\}$ we calculate the statistic $M_n^{b*} = M_n(Y_1^{b*}(\cdot), \ldots, Y_n^{b*}(\cdot))$, with M_n as given in (2.15) and denote by

$$H_{n,B}^*(x) = \frac{1}{B}\sum_{b=1}^{B} \mathbb{I}\{M_n^{b*} \le x\}$$

the empirical distribution function of $M_n^{1*}, \ldots, M_n^{B*}$. We determine the $(1-\alpha)$-quantile of this distribution and use its quantiles as critical values for the test statistic $M_n = M_n(Y_1(\cdot), \ldots, Y_n(\cdot))$. In our simulation study we made 1000 replications of this procedure with $B = 200$ bootstrap-samples, the corresponding results under the null hypothesis are presented in Table 4 for the sample sizes $n = 25$ and $n = 50$, and the regression functions (5.1) and (5.2). Compared to the test based on the normal approximation we observe a substantial improvement with respect to the approximation of the nominal level.

In Tables 5 and 6 we show the simulated rejection probabilities of the wild bootstrap test for the alternatives (5.3) and (5.4), respectively. In all cases we obtain similar rejection probabilities as for the test defined in (2.15). Compared to the test based on the asymptotic distribution, a slight loss in power is observed in case of the alternative $f_i(u,t) + \sin(2\pi t)$, while in case of the exponential alternative we observe a negligible improvement for the majority of scenarios.

As a second example we study the finite sample properties of the test for the hypothesis

$$H_0 : m(u,t) = \gamma(t) f(u,t), \tag{5.8}$$

defined in Section 3, where again $f : [0,1]^2 \to \mathbb{R}$ is some given function and $\gamma : [0,1] \to \mathbb{R}$ (i.e. $k = 1$). The discussion at the end of Section 3 suggests to reject the hypothesis H_0 if the inequality (3.6) is satisfied. We have investigated the finite sample properties of this test under the assumptions of the previous study for $f = f_1$ as given in (5.1) and $\gamma \equiv 1$. The normal approximation did not yield a sufficiently accurate approximations of the level for sample sizes up to $n = 500$ and for this

Table 6
Simulated rejection probabilities of the bootstrap test under the alternative $H_1 : m(u, t) = f_i(u, t) + \sin(2\pi t)$, $i = 1, 2$, where the regression functions f_1 and f_2 are given in (5.1) and (5.2), respectively.

	n	0.15	0.1	0.05	0.01
$f_1(u, t)$	25	0.957	0.922	0.857	0.673
	50	0.998	0.996	0.986	0.918
$f_2(u, t)$	25	0.988	0.977	0.952	0.797
	50	0.999	0.998	0.997	0.985

Table 7
Simulated rejection probabilities of the bootstrap test for the hypothesis (5.8). Under $H_0 : m(u, t) = f_1(u, t)$, under $H_1 : m(u, t) = f_1(u, t) + 1/2 \exp(t)$.

	n	0.15	0.1	0.05	0.01
H_0	25	0.148	0.106	0.062	0.022
	50	0.152	0.104	0.048	0.010
	100	0.160	0.104	0.052	0.018
	200	0.148	0.116	0.062	0.016
	500	0.150	0.104	0.06	0.012
H_1	25	0.978	0.954	0.910	0.792
	50	0.996	0.992	0.980	0.940
	100	1	1	1	0.998
	200	1	1	1	1
	500	1	1	1	1

reason these results are not depicted. As an alternative we propose to use a wild bootstrap approximation similar to the one given in the previous paragraph. More precisely, we calculate residuals analogously to (5.5) by

$$\hat{\varepsilon}(u, t_i) = Y_i(u) - \hat{\gamma}(t_i)\, f_1(u, t_i)$$

for $i = 1, \dots, n$ and $u \in [0, 1]$, where $\hat{\gamma}(t_i)$ denotes the least square estimator for $\gamma(t_i)$. As in Eqs. (5.6) and (5.7) we define

$$\varepsilon_i^{b*}(u) = v_i^{b*}\, \hat{\varepsilon}(u, t_i)$$

and

$$Y_i^{b*}(u) = \hat{\gamma}(t_i)\, f_1(u, t_i) + \varepsilon_i^{b*}(u) \qquad (i = 1, \dots, n)$$

to obtain a wild bootstrap sample. The results of the corresponding bootstrap test are shown in Table 7. We observe that the resampling procedure yields to a test with a very accurate approximation of the nominal level (again we chose the parameter $\gamma \equiv 1$) and a perfect power behavior under the alternative $H_1 : m(u, t) = f_1(u, t) + \frac{1}{2} \exp(t)$.

Acknowledgments

The authors would like to thank Martina Stein, who typed parts of this manuscript with considerable technical expertise. This work has been supported in part by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823, Teilprojekt C1) of the German Research Foundation. We are also grateful to two unknown referees for their constructive comments on an earlier version of this paper.

Appendix. Proof of Proposition 2.1

We introduce the representation

$$S_i(u) = \Delta m_{u,i} + \Delta \varepsilon_{u,i}$$

with

$$\begin{aligned} \Delta m_{u,i} &:= m(u, t_i) - m(u, t_{i-1}), \\ \Delta \varepsilon_{u,i} &:= \varepsilon(u, t_i) - \varepsilon(u, t_{i-1}), \end{aligned}$$

and consider the following decomposition of the estimate $\hat{k}$

$$\hat{k}(u, v) = 2T_{1n}(u, v) + T_{2n}(u, v) + 2T_{3n}(u, v) + \tilde{T}_n(u, v), \tag{A.1}$$

where

$$T_{1n}(u,v) = \frac{1}{4(n-3)}\sum_{i=2}^{n-2} \Delta m_{u,i}\Delta m_{u,i+2}\left\{\Delta m_{v,i+2}\Delta\varepsilon_{v,i} + \Delta m_{v,i}\Delta\varepsilon_{v,i+2}\right\}$$

$$\begin{aligned} T_{2n}(u,v) \;=\; & \frac{1}{4(n-3)}\sum_{i=2}^{n-2}\big[\Delta m_{u,i}\Delta m_{u,i+2}\Delta\varepsilon_{v,i}\Delta\varepsilon_{v,i+2} + \Delta m_{u,i}\Delta m_{v,i}\Delta\varepsilon_{u,i+2}\Delta\varepsilon_{v,i+2} \\ & + \Delta m_{u,i+2}\Delta m_{v,i+2}\Delta\varepsilon_{u,i}\Delta\varepsilon_{v,i} + \Delta m_{u,i+2}\Delta m_{v,i}\Delta\varepsilon_{u,i}\Delta\varepsilon_{v,i+2}\big] \end{aligned}$$

$$T_{3n}(u,v) = \frac{1}{4(n-3)}\sum_{i=2}^{n-2} \Delta\varepsilon_{v,i}\Delta\varepsilon_{v,i+2}\left[\Delta m_{u,i}\Delta\varepsilon_{u,i+2} + \Delta m_{u,i+2}\Delta\varepsilon_{u,i}\right]$$

$$\tilde{T}_n(u,v) = \frac{1}{4(n-3)}\sum_{i=2}^{n-2} \Delta\varepsilon_{u,i}\Delta\varepsilon_{u,i+2}\Delta\varepsilon_{v,i}\Delta\varepsilon_{v,i+2}.$$

We show that the first three terms of the decomposition (A.1) are asymptotically negligible. For this reason we analyze the term $T_{1n}(u,v)$ exemplarily. We have

$$T_{1n}(u,v) = T_{1n}^{(a)}(u,v) + T_{1n}^{(b)}(u,v) \tag{A.2}$$

with

$$T_{1n}^{(a)}(u,v) = \frac{1}{4(n-3)}\sum_{i=2}^{n-2} \Delta m_{u,i}\Delta m_{u,i+2}\Delta m_{v,i+2}\Delta\varepsilon_{v,i}$$

$$T_{1n}^{(b)}(u,v) = \frac{1}{4(n-3)}\sum_{i=2}^{n-2} \Delta m_{u,i}\Delta m_{u,i+2}\Delta m_{v,i}\Delta\varepsilon_{v,i+2}.$$

Both sums are centered and for the calculation of the variance of $T_{1n}^{(a)}(u,v)$ it follows

$$\mathrm{Var}(T_{1n}^{(a)}) = \frac{1}{16(n-3)^2}\sum_{i=2}^{n-2}\sum_{j=2}^{n-2} E\left[\Delta m_{u,i}\Delta m_{u,i+2}\Delta m_{u,j}\Delta m_{u,j+2}\Delta m_{v,j+2}\Delta m_{v,i+2}\Delta\varepsilon_{v,i}\Delta\varepsilon_{v,j}\right].$$

Note that this sum is dominated by the sum of those expectations corresponding to the indices with $i = j$, $i = j+1$ or $j = i+1$. We exemplarily treat the case $i = j$. Using the Lipschitz continuity of the function m it follows

$$\Delta m_{u,i+2} \le \max_{2\le i\le n} |t_i - t_{i-1}|^{\gamma} = O(n^{-\gamma}) \tag{A.3}$$

uniformly with respect to $i = 2, \ldots, n$, and this estimate yields

$$E[\Delta^2 m_{u,i}\Delta^2 m_{u,i+2}\Delta^2 m_{v,i+2}\Delta^2\varepsilon_{v,i}] = O\left(n^{-6\gamma}\right).$$

Consequently, by Markov's inequality we obtain (uniformly with respect to u and v)

$$T_{1n}^{(a)}(u,v) = O_p(n^{-3\gamma}).$$

The term $T_{1n}^{(b)}(u,v)$ in (A.2) is treated similarly, which implies $T_{1n}(u,v) = o_p\left(n^{-1/2}\right)$. Similar arguments for the statistics $T_{2n}(u,v)$ and $T_{3n}(u,v)$ in (A.1) give

$$\hat{k}(u,v) = \tilde{T}_n(u,v) + o_p(n^{-1/2}).$$

For the investigation of the remaining (dominating) term $\tilde{T}_n(u,v)$ we note that the sequence $(\Delta\varepsilon(u,t_i))_{i=1,..n}$ is 2-dependent, which yields

$$\begin{aligned} E[\Delta\varepsilon_{u,i}\Delta\varepsilon_{u,i+2}\Delta\varepsilon_{v,i}\Delta\varepsilon_{v,i+2}] &= E[\varepsilon(u,t_i)\varepsilon(v,t_i) + \varepsilon(u,t_{i-1})\varepsilon(v,t_{i-1})] \\ &\quad \times E[\varepsilon(u,t_{i+2})\varepsilon(v,t_{i+2}) + \varepsilon(u,t_{i+1})\varepsilon(v,t_{i+1})] \\ &= [r(t_i,u,v) + r(t_{i-1},u,v)][r(t_{i+2},u,v) + r(t_{i+1},u,v)] \\ &= 4r(t_i,u,v)r(t_{i+2},u,v) + O(n^{-\gamma}) \end{aligned}$$

by the Lipschitz continuity of the covariance function r. Observing the definition of $\tilde{T}_n(u,v)$ this gives

$$E[\tilde{T}_n(u,v)] = \frac{1}{n-3}\sum_{i=2}^{n-2} r(t_i,u,v)r(t_{i+2},u,v) + O(n^{-\gamma}) = \int_0^1 r^2(t,u,v)h(t)\mathrm{d}t + o\left(n^{-1/2}\right).$$

A similar calculation shows that the variance of $\tilde{T}_n$ is of order $O(n^{-1})$, which yields the assertion of Proposition 2.1. □

References

[1] N.I. Achieser, Theory of Approximation, Frederick Ungar Publishing Co., New York, 1956.
[2] P. Besse, J.O. Ramsay, Principal components of sampled functions, Psychometrika 51 (1986) 285–311.
[3] P. Bickel, M.J. Wichura, Convergence criteria for multiparameter stochastic processes and some applications, Annals of Mathematical Statistics 42 (1971) 1656–1670.
[4] P. Billingsley, Convergence of Probability Measures, Wiley, New York, 1968.
[5] D. Bosq, Linear Processes in Function Spaces, in: Lecture Notes in Statistics, Springer, New York, 2008.
[6] F. Brodeau, Tests for the choice of approximative models in nonlinear regression when the variance is unknown, Statistics 24 (2) (1993) 95–106.
[7] H. Cardot, F. Ferraty, A. Mas, P. Sarda, Testing hypotheses in the functional linear model, Scandinavian Journal of Statistics 30 (2003) 241–255.
[8] H. Cardot, A. Goia, P. Sarda, Testing for no effect in functional linear regression models, some computational approaches, Communications in Statistics – Simulation and Computation 33 (2004) 179–199.
[9] A. Cuevas, M. Febrero, R. Fraiman, Linear functional regression: the case of fixed design and functional response, Canadian Journal of Statistics 30 (2002) 285–300.
[10] I. Delsol, Régression sur variable fonctionnelle: Estimation, Tests de structure et Applications, http://thesesups.ups-tlse.fr/399/1/DelsolLaurent.pdf, Ph.D. thesis, Université Paul Sabatier Toulouse, 2008.
[11] H. Dette, S. Biedermann, Testing linearity of regression models with dependent errors by kernel based methods, Test 7 (2000) 417–438.
[12] H. Dette, A. Munk, T. Wagner, Testing model assumptions in multivariate linear regression models, Nonparametric Statistics 12 (1999) 309–342.
[13] M. Escabias, A.M. Aguilera, M.J. Valderrama, Principal component estimation of functional logistic regression: discussion of two different approaches, Journal of Nonparametric Statistics 16 (2004) 365–384.
[14] J.J. Faraway, Regression analysis for a functional response, Technometrics 39 (1997) 254–261.
[15] F. Ferraty, P. Vieu, Curves discrimination: a nonparametric functional approach, Computational Statistics and Data Analysis 44 (2003) 161–173.
[16] F. Ferraty, P. Vieu, Nonparametric Functional Data Analysis: Theory and Practice, Springer, New York, 2006.
[17] A.R. Gallant, Nonlinear Statistical Models, Wiley, New York, 1987.
[18] W. González Manteiga, J. Vilar Fernández, Testing linear regression models using non-parametric regression estimators when the errors are correlated, Computational Statistics and Data Analysis 20 (1995) 521–541.
[19] D. Hlubinka, L. Prchal, Changes in atmospheric radiation from the statistical point of view, Computational Statistics and Data Analysis 51 (2007) 4926–4941.
[20] A. Kneip, K. Utikal, Time trends in the joint distribution of income and age, in: Economic Essays. A Festschrift for Werner HILDENBRAND, Springer Verlag, 2001.
[21] P. Kokoszka, I. Maslova, J. Sojka, L. Zhu, Testing for lack of dependence in the functional linear model, The Canadian Journal of Statistics 36 (2008) 207–222.
[22] A. Mas, Testing for the mean of random curves: a penalization approach, Statistical Inference for Stochastic Processes 10 (2007) 147–163.
[23] H.G. Müller, U. Stadtmüller, Generalized functional linear models, Annals of Statistics 33 (2005) 774–805.
[24] S. Orey, A central limit theorem for m-dependent random variables, Duke Mathematical Journal 52 (1958) 543–546.
[25] J.O. Ramsay, C.J. Dalzell, Some tools for functional data analysis (with discussion), Journal of the Royal Statistical Society, Series B 52 (1991) 539–572.
[26] J.O. Ramsay, B.W. Silverman, Functional Data Analysis, second ed. Springer, New York, 2005.
[27] J. Sacks, D. Ylvisacker, Designs for regression problems with correlated errors III, Annals of Mathematical Statistics 41 (1970) 2057–2074.
[28] G.A.F. Seber, C.J. Wild, Nonlinear Regression, John Wiley and Sons Inc., New York, 1989.
[29] Q. Shen, J. Faraway, An F test for linear models with functional responses, Statistica Sinica 14 (2004) 1239–1257.
[30] C.F.J. Wu, Jackknife, bootstrap and other resampling methods in regression analysis, Annals of Statistics 14 (1986) 1261–1295.
[31] X. Yang, Q. Shen, H. Xu, S. Shoptaw, Functional regression analysis using an F test for longitudinal data with large numbers of repeated measures, Statistics in Medicine 26 (2007) 1552–1566.